DOCUMENT RESUME

ED 104 844 SP 009 080

AUTHOR Borg, Walter R.
TITLE Making the Leap from Correlational to Experimental Studies of Teacher Behavior.
PUB DATE 30 Mar 75
NOTE 16p.; Paper presented at the Annual Meeting of the American Educational Research Association (Washington, D.C., March 30, 1975)

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS *Academic Achievement; Behavior Change; Classroom Techniques; *Correlation; Educational Research; Experimental Teaching; Higher Education; *Research; Student Teacher Relationship; *Teacher Behavior; Teacher Education; *Teacher Evaluation; Teaching Skills
IDENTIFIERS *Utah State University

ABSTRACT

What little we know about relationships between specific classroom behavior of teachers and relevant pupil outcomes has been obtained almost entirely from coreelational studies. Yet if we are ever to apply knowledge in this area to teacher education, we must carry out experimental studies in which teachers are trained to emit specific behaviors that are found to be related to pupil outcomes. In order to evaluate how specific changes in teacher behavior can bring about changes in related pupil outcomes, Utah State University created three sets of protocol modules employed as experimental treatments. Through these studies, it was determined that relationships between specific teacher behaviors and pupil outcomes tend to be higher in correlational studies than in experimental studies. Four variables seem to be the cause of this difference. First, general teaching competence operates more strongly in correlational studies and probably leads to spuriously high correlations between specific teacher behaviors and pupil outcomes. Second, the length of pupil exposure to teaching behaviors studied may be longer in correlational studies than in experimental studies. Third, when teachers adopt new behaviors, there is a lag in the development by pupils of appropriate responses to these behaviors. Fourth, when teachers are trained in new behaviors, time is not often allowed to incorporate the skill into teacher performance. (Author/JS)

ED104844

AERA Session 3.03

Making the Leap from Correlational to Experimental Studies of Teacher Behavior[1]

by

Walter R. Borg

UTAH STATE UNIVERSITY

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

What little we know about relationships between specific classroom behavior of teachers and relevant pupil outcomes has been obtained almost entirely from correlational studies. For example, Rosenshine's recent review of studies of teacher behavior and student achievement deals almost entirely with correlational research (1971). He points out that there have been few experimental studies done in this area in which teachers have been trained in a specified manner, measures have been made of their instructional behavior and student achievement has been compared in the classrooms of experimental vs. control teachers. Yet, it seems apparent that if we are ever to apply our knowledge in this area to the training of teachers, we must carry out experimental studies in which teachers are trained to emit specific behaviors that are found to be related to pupil outcomes.

Probably a major factor that has kept researchers from moving into experimental studies in this area is that very few adequate experimental treatments have been available. In other words, it has only been recently that teacher training materials and strategies have been developed which focus on specific teacher behavior and which have been demonstrated to bring about significant changes in the way that teachers perform in the classroom. For example, the minicourses developed at the Far West Laboratory have all been evaluated using pre-post designs in which the teacher's classroom behavior has been measured before and after training in order

SP009 080

[1]This paper was presented at the annual meeting of the American Educational Research Association, Washington, D.C., March 30, 1975.

to determine whether the training brings about significant behavioral changes (Borg, et al, 1970). Some minicourse evaluations have used control group designs while others have employed single group designs. For both kinds of evaluation the results have generally indicated that the minicourses do bring about significant changes in the classroom behavior of teachers, and could thus be employed as experimental treatments in studies concerned with the degree to which specific changes in teacher behavior bring about changes in related pupil outcomes. Some of the minicourse evaluations have taken this next step and measured changes in pupil outcomes between pre and post observations and/or between the pupils of experimental and control group teachers although this has not been the emphasis in most of these evaluations.

During the past two years I have been involved in similar evaluations in which the three sets of Utah State University Protocol Modules have been employed as experimental treatments. In developing these modules, we drew heavily upon previous correlational studies as a source of ideas on specific teacher variables that appear to be related to pupil outcomes. We have been concerned not only with bringing about changes in teacher behavior but also have given much attention to linking these changes to pupil outcome variables. In the process of carrying out five such studies over the past two years, I have become increasingly aware of a number of problems that occur when an investigator, building upon the evidence from correlational research, attempts to change teacher behavior and related pupil outcomes in experimental studies. The purpose of this paper is to discuss the probable causes behind one major problem we have encountered--the tendency for experimental studies to yield lower relationships between specific teaching behaviors and pupil outcomes than would be expected from our knowledge of the correlational data.

A General Teaching Factor?

In correlational studies conducted with experienced teachers, the teacher behaviors studied have probably been learned by the more competent teachers over a period of years. Observations and interviews I have carried out with effective teachers over the past few years have suggested that they have developed their competence through the process of trying new approaches and utilizing pupil feedback in order to decide whether each approach should be continued or not. In many cases, this is probably not a conscious process but is simply a matter of the teacher continuing those activities which bring reinforcement in the form of desired pupil responses and discontinuing activities that do not bring about reinforcement. The critical elements in this process are probably (1) the flexibility of the teacher, that is, the capacity of the teacher to think of and try new teaching behaviors and (2) the teacher's sensitivity to pupil feedback. A teacher who tries a great many teaching behaviors and is very perceptive of pupil feedback should develop a range of effective teaching techniques on a trial and error basis much more rapidly than a teacher who rarely tries anything new and is insensitive to the effects of new teaching behaviors upon his or her pupils. If these two teacher characteristics are in fact operating and account for some of the performance differences between effective and ineffective teachers, then we could expect a general competence factor to emerge in that some teachers would be using a wide range of specific behaviors more effectively than other teachers.

In correlational research, however, this situation could lead to obtaining spuriously high correlations between the specific teacher behaviors that the investigator chooses to study and pupil outcomes such as achievement. Let us suppose, for example, that an investigator

selects six specific behaviors and carries out a correlational study to determine whether these behaviors are significantly related to pupil achievement. Typically, his sample would consist of teachers who differ along a continuum on the two characteristics of flexibility and sensitivity to pupil feedback. He would observe the six selected behaviors in the classrooms of teachers making up his sample and would also administer an achievement measure to their pupils. He would then compute correlat probably partialling out pupil ability in order to determine the significance of relationships between use of each of the six teaching behaviors and pupil achievement. However, if my hypothesis concerning flexibility and feedback sensitivity is correct, teachers who use the six behaviors that the investigator selected would probably also use a wide range of other specific teaching behaviors that he had not identified or observed but which also relate to pupil achievement. Therefore, teachers who are effective on these six behaviors are likely to be effective on a great many other behaviors and conversely, teachers who are weak on the six behaviors are likely to be weak on a great many other behaviors related to pupil achievement. The result could be that the six behaviors the investigator selected could appear to be more highly related to pupil achievement than they really are because of the fact that they tend to occur or not occur in conjunction with a great many other equally important behaviors.

Let us suppose that we now take the results of this correlational study and set up an experimental study in which we train a group of teachers, none of whom makes extensive use of the six behaviors prior to training. After training, our experimental teachers would be emitting these six behaviors at a much higher rate than before training. If we then correlate the teachers' use of these six behaviors with pupil

achievement on a content unit taught after the teachers have been trained, we would probably obtain much lower correlations than those that had been obtained in the correlational study. This is because teacher performance on these six variables in the experimental study is due largely to training and therefore is not related to overall teacher competence to the same degree that it would be in the correlational study.

To test this hypothesis concerning the difference in relationships that can emerge from correlational as opposed to experimental studies I decided to compare correlations obtained in one of the better correlational studies (Wright and Nuthall, 1970) with those obtained in one of the Utah State University experimental studies completed earlier this year (Borg, 1975). In the Wright and Nuthall study, the teacher group consisted of six highly experienced teachers, five student teachers who had completed a two-year teacher training course, and six student teachers who were in the first year of teacher training and had had no previous practice teaching experience. All of the teacher sample taught third grade classes in the city of Christchurch, New Zealand. The investigators developed three ten-minute lessons dealing with the black-backed gull and made up a 25-item multiple choice achievement test aimed primarily at factual information covered in the lessons. The participating teachers were supplied with two stuffed and mounted specimens of black-backed gulls, reproductions of material taken from two standard texts on New Zealand birds and an outline of content for the three lessons. They were instructed to teach this content only during the designated times. Audiotape recordings were made of the three ten-minute lessons and on the day following the third lesson pupils were given the achievement test. Residual achievement test scores were computed using multiple regression analysis in order to eliminate the influence of pupil intelli-

gence and pupil knowledge of related natural science concepts. Correlations between the residual achievement score and teacher use of each classroom behavior were then computed. Five of the teacher behavior variables observed by Wright and Nuthall were very similar to five of the variables observed in the Utah State University study.

In the Utah State University study a total of 40 elementary teachers participated. These teachers taught grades four, five and six. Observation of the teachers' use of twelve specific teaching behaviors was made at the beginning and end of the study. The 40 participating teachers were divided into an experimental group of 25 teachers who were trained on specific teacher behaviors over a period of five weeks and a control group of 15 comparable teachers who received no training. At the completion of the training period, all 40 teachers were asked to teach a content unit based on a National Geographic article which dealt with aging and focused especially on three areas in the world where an unusual proportion of the population live beyond the age of 100. The teaching unit given to the teachers included a copy of the National Geographic article, ten 35mm color slides related to the article and a simplified reading lesson that could be used in conjunction with the article. Teachers taught the content of the unit 50 minutes per day for four consecutive days. A 40-item multiple choice achievement test dealing only with content related to the unit was administered on the fifth day. Each teacher was observed for two of the four days during which the lesson was taught. These days were selected at random and no teacher knew which two days he or she would be observed. In addition to other analysis, partial correlations were computed between the use of each specific teacher behavior by the 40 teachers and achievement of their pupils. Pupil ability, socio-economic status, and a measure of the degree to which

each teacher covered the main concepts of the unit were partialled out.

Although these two studies were far from identical, they are similar to one another in many ways. Both used special content units not part of the regular curriculum and developed multiple choice tests from these units. Both were carried out at the elementary school level, although not at the same grade level. Both adjusted correlations between pupil achievement and teacher behavior for differences in pupil ability. Both studies dealt with five teacher variables that were very similar although not identical. A major difference in the studies is that in the U.S.U. study, the experimental group of teachers were trained to emit the specific teacher behaviors in the classroom, and in fact there were significant changes in the teachers' use of each of these five behaviors between the pre and post-training observations. The U.S.U. sample, however, also included 15 control group teachers who had not been trained.

Since teachers in the Wright and Nuthall study had not been trained in the specific behaviors, their performance can be considered primarily a result of an accumulation of skills gained through differing amounts of trial and error experience in the classroom. In contrast, the behavior of teachers in the U.S.U. study would represent a combination of trial and error learning that occurred prior to the study plus specific training, for the experimental teachers, in the behaviors. Also, if our hypothesis is valid, the more effective teachers in the Wright and Nuthall study would be more likely because of greater flexibility and feedback sensitivity, to have developed a related set of skills that were not measured in the research. Thus, the correlations between specific teacher behaviors and pupil achievement should be higher in the Wright and Nuthall study than in the U.S.U. study. A comparison showed that four of the five correlations obtained in the Wright and Nuthall study were

higher than those obtained in the U.S.U. study.

After completing this comparison, it occurred to me that a much more precise test of my hypothesis could be made by separately computing partial correlations in the U.S.U. study for the experimental and control group teachers. We hypothesized that the control group correlations would be higher since teachers in the control group who were using the behaviors had probably learned these along with other related behaviors over a period of time because of greater flexibility and sensitivity to pupil feedback. In contrast, the performance of the experimental group teachers would be a combination of previously learned behaviors plus the results of the rather intensive training that they underwent during the five weeks in the experimental study which focused only on twelve specific behaviors. In other words, high performance by control group teachers would imply not only an ability to use the behaviors that were being observed but also a general competence to use a variety of related behaviors. This would be less true in the case of the experimental group teachers whose high performance on the trained behaviors would not necessarily indicate a comparable degree of competence in related behaviors that had not been covered in the training.

It may be seen in Table 1 that the data supported our hypothesis for all 10 of the positive teacher behaviors that had been covered in the training. Many of the differences between our two groups are quite large and suggest that the general teaching factor we hypothesized could be of major importance in producing spuriously high relationships in correlational studies in which specific teacher behaviors are correlated with pupil outcomes.

It is difficult to explain why the pattern of differences found for the ten positive teacher behaviors did not maintain for the two negative

behaviors. Some of our evaluation data from Minicourse I suggested that it is easier for teachers to extinguish a negative behavior than to add a positive behavior to their teaching repertoire. Changes that teachers made in the negative behaviors in Minicourse I also appeared to be more permanent than changes in positive behaviors when we observed their teaching three years after training (Borg, 1972). The process that teachers employ in extinguishing an undesired behavior is somewhat different than the process used to add a desired behavior to the teacher's repertoire. But I cannot see how this difference could have led to the higher negative correlations obtained for our experimental group teachers in the U.S.U. protocol evaluation study. One major difference between negative and positive teacher behaviors is that the skill component is not a factor in negative behaviors. A new positive behavior such as making specific praise statements may be used more or less skillfully but a negative behavior such as asking multiple questions seems unrelated to skill. Is it possible for the teacher to NOT ask a multiple question in a more or less skillful manner?

The teacher who tries new teaching behaviors and is sensitive to pupil feedback would probably learn to avoid multiple questions since they obviously confuse many pupils. But there seems no reason to believe that teachers who are not flexible and sensitive would make extensive use of this negative behavior.

It would seem logical to hypothesize that negative behaviors would correlate about the same with pupil outcomes in experimental and correlational studies. Since this was not the case in our study (Borg, 1975) I plan to explore this question further, using data from another study we have recently completed. However, if the correlations between negative teacher behaviors and pupil outcomes in this later study also turn

out to be higher under the experimental research condition we will have to look for other theoretical constructs to explain this result.

Although general teaching competence appears to be the most important factor leading to differences between the outcomes of experimental and correlational studies, there are some other factors operating that also tend to help produce weaker results in experimental studies.

Length of Pupil Exposure

For example, in correlational studies of experienced teachers, we have suggested that the teachers who are emitting the behaviors being studied probably have developed skill in using these behaviors over a long period of time. This would mean in turn that their pupils would have been exposed to these behaviors during the entire time they have been in the teacher's classroom. Therefore, in correlational studies pupil outcome measures administered near the end of the school year would be influenced by the cumulative effect of the specific teaching behaviors being studied over a relatively long period of time. On the other hand, in experimental studies in which teachers are trained to emit specific behaviors, the pupils are usually exposed to the behaviors for a much shorter period of time. Even if the experimental treatment starts early in the school year, pupils will not be exposed to any great extent to the new behaviors the teacher has learned until the training has been completed. In the interim, the students may be exposed to behaviors that are in conflict with those the teacher will learn during the training.

Pupil Performance Lag

A related problem found in experimental studies is that when teachers adopt new teaching behaviors after training in an experimental treatment, it may be some time before their pupils adjust to these new

teaching behaviors. For example, one of the objectives of Minicourse I was to decrease the proportion of fact questions asked by elementary school teachers and to increase the proportion of higher-order questions. In observing videotapes that we made in teachers' classrooms after they had completed Minicourse I, I noticed that pupils frequently attempted to give fact answers to higher order questions. It seems likely that these pupils had been exposed in the past to such a high proportion of fact questions that they had developed a successful procedure for coping with teacher questions; specifically, by giving fact answers. When the teacher suddenly began using a great many higher-order questions, many children seemed to have difficulty coping with this new situation and persisted in the fact-giving behavior that had been successful in the past. This tendency for adjustments in pupil behavior to lag somewhat behind changes in teacher behavior is a factor that could have an important effect on pupil outcomes obtained in experimental studies. This effect could be particularly strong in studies in which pupil outcomes are measured very shortly after training of teachers in new behaviors has been completed.

Incorporating New Behaviors Into Teaching

Another factor that can reduce the significance of relationships between teacher performance and pupil outcomes in experimental studies is that teachers in such studies who have just learned new teaching skills or behaviors usually have not yet incorporated these behaviors into their natural mode of teaching. While I was trianing teachers in the skills covered in the U.S.U. Protocol Modules, I frequently had teachers mention that they felt unnatural or uneasy in using the new behaviors they have learned. As is the case with most skills, a person learning a new teaching skill is likely to use this skill clumsily or

ineffectively during the learning stage. It takes considerable classroom practice and feedback before a teacher is comfortable with the use of a new skill and has learned when the skill best fits into the teaching situation. Therefore, in experimental studies, the training should not stop until teachers have mastered the skills to the point where they use them naturally and appropriately in the teaching situation. Most experimental studies that have been done to date involving the training of teachers in specific skills have carried out the post-evaluation very shortly after the training has been completed. In most cases, these studies (such as the Minicourse evaluations) have not provided sufficient time for participating teachers to fully master the skills being learned.

In contrast, teachers who participate in correlational studies have probably developed whatever teaching behaviors they have over several years and therefore, those who use the behaviors being studied are likely to be more skillful in applying them in the classroom.

Summary

In summary, we have observed that relationships between specific teacher behaviors and pupil outcomes tend to be higher in correlational studies than in experimental studies. We have identified four variables that appear to be probable causes of this difference. The first of these is that general teaching competence operates more strongly in correlational studies and probably leads to spuriously high correlations between specific teacher behaviors and pupil outcomes. Second, the length of pupil exposure to the teaching behaviors being studied is likely to be longer in correlational studies than in experimental studies. Third, when teachers adopt new behaviors (as they do when given training in an experimental treatment), there is a lag in the development by pupils of appropriate responses to these new teacher behaviors. Thus,

if pupil outcomes are measured too soon after teacher training, spuriously low results will be obtained.

Fourth, when teachers are trained in new behaviors, a time elapses after they have partially learned the behaviors and can emit them in the classroom before they can use them skillfully. If experimental treatments do not allow sufficient time for teachers to develop skill in using newly learned behaviors, correlations between the partially mastered behaviors and pupil outcomes are likely to be lower than those obtained in correlational studies in which teachers are more likely to have fully mastered the behaviors being observed.

Table 1

Correlations between specific teacher behaviors and pupil achievement in the USU Protocol evaluation

Teacher Behavior	Experimental Teachers (N=25) M	SD	r	Control Teachers (N=15) M	SD	r
1. Multiple Questions*	5.15	4.78	-.43	3.77	5.13	-.24
2. Cueing	36.1	20.3	.28	31.3	22.9	.75
3. Defining	27.4	7.8	.31	16.0	8.7	.57
4. General Praise	96.0	62.3	.46	50.3	43.6	.59
5. Specific Praise	21.1	11.4	.03	6.5	12.6	.66
6. Paraphrasing	15.1	8.5	.31	12.7	14.4	.60
7. Use of Student Ideas	13.1	7.2	-.04	7.0	5.7	.25
8. Voice Modulation	17.0	11.3	.32	13.6	10.3	.70
9. Terminal Structure	7.9	8.3	.37	5.8	8.1	.39
10. Vague Words*	13.9	7.3	-.30	15.4	8.6	-.05
11. Summary Review	6.4	1.3	-.04	5.6	2.4	.15
12. Opening Review	6.4	1.4	.26	4.5	2.0	.30

*These are negative behaviors, i.e. the aim of the experimental treatment was to reduce teachers' use of the behavior.

References

Borg, W. R. The minicourse as a vehicle for changing teacher behavior: a three-year followup. Journal of Educational Psychology, 1972, 63 (6), 572-579.

Borg, W. R., Kelley, M. L., Langer, P., & Gall, M. The minicourse: a microteaching approach to teacher education. Beverly Hills, Calif.: Macmillan Educational Services, Inc., 1970.

Borg, W. R. Protocol materials as related to teacher performance and pupil achievement. Manuscript submitted for publication, 1975.

Rosenshine, B. Teaching behaviours and student achievement. London: National Foundation for Educational Research in England and Wales, 1971.

Wright, C. J., & Nuthall, G. Relationships between teacher behaviors and pupil achievement in three experimental elementary science lessons. Americal Educational Research Journal, Nov. 1970, 7 (4), 477-491.